Hernandez-Carrion, J. R. and González-Rodriguez, D., "Modelling Complex Dynamics and Distributed Generation of Knowledge with Bacterial-Based Algorithms" in *Economics and Bussiness Communication Challenges. International Week*, W. Dyduch. M. Pankowska (eds.). Katowice: University of Economics in Katowice, 2014, pp.138–148. ISBN: 978-83-7875-220-2

# Modeling Complex Dynamics and Distributed Generation of Knowledge with Bacterial-based Algorithms

**Jose Rodolfo Hernandez-Carrion[1] and Diego Gonzalez-Rodriguez[2]**

[1] University of Valencia, Department of Applied Economics, Spain, rodolfo@uv.es
[2] Universidad Carlos III, Madrid, Spain, xmunch@xmunch.com

## Abstract

This study aimed to test that connected and heterogeneous societies with peer-to-peer (P2P) exchanges are more resilient than centralized and homogeneous ones. In agent-based modeling, agents with bounded rationality interact in a common environment guided by local rules, leading to Complex Adaptive Systems that are named 'artificial societies'. These simplified models of human societies grow from the bottom up in computational environments and can be used as a laboratory to test some hypotheses. We have demonstrated that in a model based on free interactions among autonomous agents, optimal results emerge by incrementing diversity and decentralization of communication structures, as much as in real societies Internet is leading to the emergence of improvements in collective intelligence. In order to achieve a real "Knowledge Society", what we have named a "P2P Society", it is necessary to increase decentralization and heterogeneity through information policies, distributed communication networks, open e-learning approaches and initiatives like public domain licenses, free software and open data.

**Key words:** Complex Adaptive Systems, P2P Society, Collective Intelligence, Bacterial-based Algorithms

## 1 Introduction

Resilience is the capacity of a system to absorb changes in environment, adapting its properties to disturbance but retaining its basic structure (Deffuant & Gilbert, 2011). Our paper defends that connected and diverse societies with *peer-to-peer or P2P* (Bauwens et al., 2005) exchanges are more resilient than centralized and homogeneous ones. We consider that the increase of *P2P* communications through the Internet, by crossing cultural boundaries without constraints, establish an inflection point in human evolution. The resulting *P2P Society* will be more organic, more efficient and more adaptable to changes than older social systems, leading to an important shift in different areas such as business, education or governance.

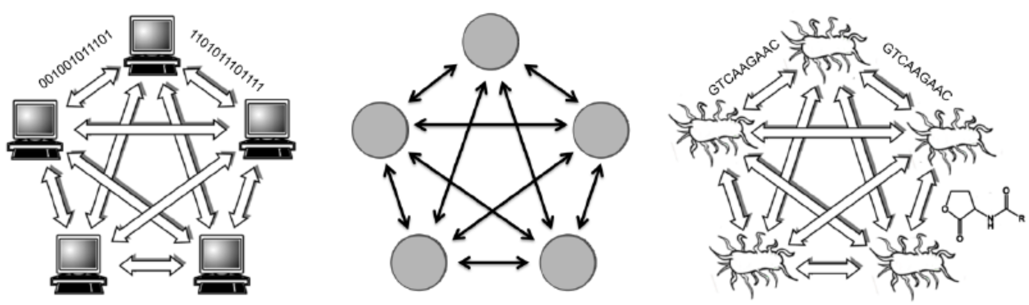


Figure 1. In peer-to-peer (P2P) networks there is not any node with central control, they follow a distributed network architecture. Furthermore, all the nodes or peers act as consumers and providers of resources. From an information-processing perspective, peers are able to exchange data leading to decentralized computation processes.

This issue has a special interest when we focus on education, because P2P dynamics have supposed a breakthrough in the way we consume and produce information. In a *P2P Society*, citizens exchange data without intermediaries. Rather than before, when learning processes implied passive information reception from centralized sources, nowadays education can be participatory and distributed. Digital networks constitute bridges that allow different individuals to work together even though they are located far away from each other. But collective generation of knowledge does not depend only on connectivity; it is also linked to diversity. In a plural world, everyone can provide useful value to the global conversation. For example, multidisciplinary approaches in research require professionals with different backgrounds in order to produce innovation. A computer scientist and a lawyer can pursue a common goal and develop a solution together. So when we talk about a *P2P society*, we talk about putting together apparently unrelated seeds in order to produce something new and better (Bauwens, 2005).

We sustain that social systems mainly need decentralization and heterogeneity to develop an optimal scenario and become more resilient. In order to demonstrate this hypothesis, we have been looking for an appropriate paradigm to model artificial societies (Cioffi-Revilla & Rouleau, 2010; Mitchell, 1999; Tesfatsion, 2003). Finally, after studying how information exchanges occurs in nature (Barabási & Oltvai, 2004), we have chosen a new approach based on bacterial conjugation, that is, a distributed communication system used by bacteria to exchange strategies of survival implemented on genetic code.

Bacterial conjugation matches the kind of dynamics we want to model because of several reasons. First of all, it is because we conceive societies as *Complex Adaptive Systems (CAS)* (Lansing, 2003) which evolution depends on interactions among autonomous agents. Secondly, because we sustain our thesis on decentralized communications (Baran, 1964), on production of knowledge in a distributed way by using *P2P* networks to share codified blocks of information (Fig. 1). Third, because even though communication and *P2P* dynamics play an important role, also heterogeneity in population and variation (or mutation) of strategies is a factor of evolution.

Bacteria have demonstrated an amazing capacity to overcome environmental changes by collective adaptation through genetic exchanges. By using a distributed communication system and sharing their individual strategies, bacteria propagate mutations as innovations that allow them to survive in different environments (González Rodríguez, 2011). Similarly, we consider that a “P2P Society” would be more resilient than a centralized and homogeneous one so by using bacterial-based algorithms we can support this idea with experimental data. In this paper we will introduce our bacterial-based approach to model artificial societies and we will test some of our ideas about distributed production of knowledge and emergence of collective intelligence. We will demonstrate that in artificial societies based on interactions among agents with bounded rationality, optimal results emerge by incrementing heterogeneity levels and decentralization of communication structures (Heylighen, 1999).

# 2 Model

## 2.1 Definition

Following an agent-based modeling approach, we want to simulate and analyze the impact of both *peer-to-peer* connections and heterogeneity on strategies optimization, that is, on distributed generation of knowledge.

In agent-based modeling, agents with bounded rationality interact in a common environment guided by local rules, leading to *Complex Adaptive Systems* that are named 'artificial societies' (Epstein & Axtell, 1996). These simplified models of human societies grow from the bottom up in computational environments and can be used as a laboratory to test some hypothesis. In our case, we will use a special type of agent-based model, a bacterial-based algorithm that is inspired by bacterial conjugation and that matches with our purpose of simulate the emergence of collective intelligence.

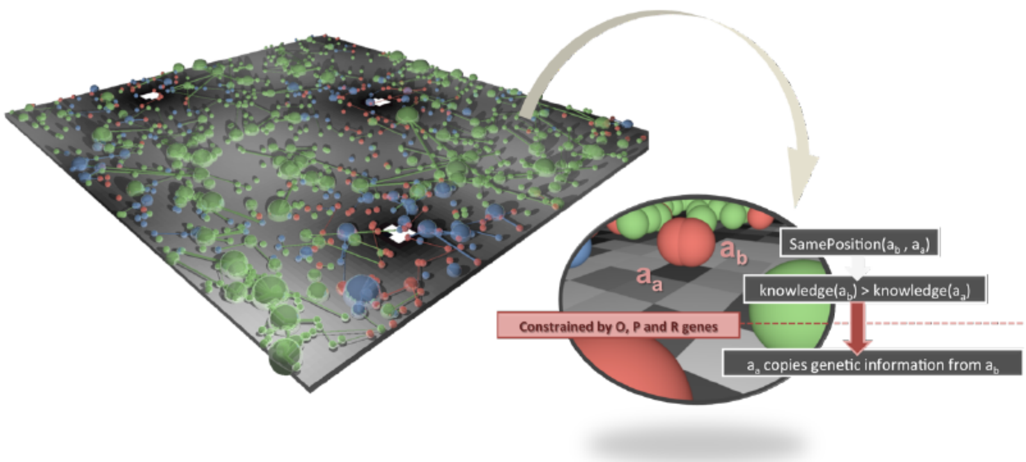


Figure 2. Comparing knowledge levels in a common cell of the bi-dimensional grid.

In this model, we have a set A with N agents ($a_i$) that represent human actors. Each agent owns a genome that contains a specific strategy ($s_i$) to optimize a function. Depending of agent's strategy, his knowledge level will be greater or lower. Then if an agent is able to optimize a given function in order to get a result with 70% of accuracy by using his strategy, his knowledge level will be set to 70 and so on. Knowledge levels determine agents' position in social structure. So agents with more successful genome will dominate cultural life of society.

During each interaction of simulation, agents move randomly through a bi-dimensional grid (Fig. 2). When two agents have the same coordinates (x,y) they meet to each other and compare their knowledge levels. After that, the one with a lower knowledge ($a_a$) want to learn from the more successful ($a_b$) so he tries to get a copy of his genome. But the agent who owns the best strategy ($a_b$) is who decides if genome is going to be shared or not because he is the owner. If the owner ($a_b$) does not share his strategic knowledge we will say that conjugative machinery to send plasmids is inhibited. Otherwise he will offer a plasmid with a copy of his genome to agents in the same coordinates and lower knowledge. Even though if the owner ($a_b$) allows the other agent ($a_a$) to get a copy of his genome and then improve his strategic knowledge he also can impose two restriction policies to that copy inhibiting:

a) *Reproduction:* The receiver of a plasmid ($a_a$) is allowed to use the strategy that is contained in the copy but he does not own the intellectual property of the strategy. Then plasmid cannot be sent to others once it is received. In this case the first owner ($a_b$) is the only one with reproduction rights on his strategy.

b) *Mutation:* The receiver ($a_a$) can use the strategy but he cannot modify it. Genome only can be used as a unit of privative software or as a behavioral dogma, following the exact strategy proposed by the agent who invented it ($a_b$). Otherwise, if mutation is not inhibited, strategies may be modified or mixed with other ones by the receiver ($a_a$).

With this model we want to show that centralized and homogeneous societies, those with greater number of agents that follow restrictive behaviors, lead to lower levels of knowledge and higher levels of inequality than distributed and heterogeneous ones. We will do it but comparing bacterial-based societies with different configurations and observing how inhibiting plasmid conjugation, reproduction or mutation modifies the statistical results. Only in a “P2P Society”, by sharing individual information among agents without communication constraints, optimal strategies and social development are achieved.

### 2.2 Agent Genome

Each agent ($a_i$) of the agents set A has its own strategy ($s_i$) coded as a part of its genome. Considering a set Sec containing several strategies ($s_i$), its cardinality |Sec| will be equal or bigger than unity and equal or smaller than cardinality of A. If by default the value of |Sec| was one, simulation would start in a completely homogeneous society. If this value was near to |A| it would be a heterogeneous society. Genome also can include another three sequences (P, R and O) that are related to the three constraints that we have described: inhibit mutation (O), inhibit original plasmid conjugation (P) and inhibit copy reproduction (R). The expression probability of these genes (OPR = O-Prob, P-Prob, R-Prob) will change the structure of the system (Fig. 3).

If there is P then the genome will not be released by conjugation, that is, that strategy will be private. So only the absence of P enables the first owner of the genome to act as a donor, that is, to send a copy of genome as a plasmid to another agent by using conjugation. If possibility of that P occurs is high then the society will follow a centralized paradigm, that is, just some nodes will be able to send information. P implies that original genome will never be copied and sent to anybody else. Then, strategies of nodes without P and a successful strategy coded on S will dominate the culture.

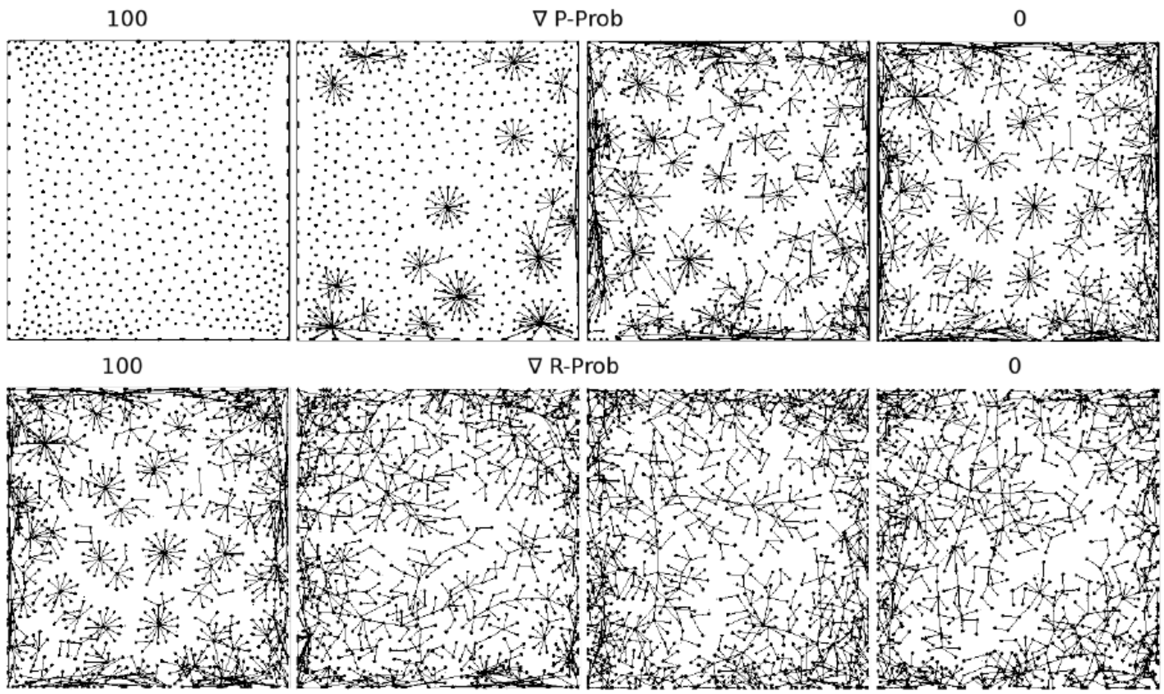


Figure 3. Impact of P-Prob and R-Prob on social structure. In the top image, the concentration of P genes in population decreases from 100% to 50%, 25% and finally 0%. The social structure changes, from unconnected nodes (P-Prob = 1) to centralized networks (P-Prob = 0.5) to decentralized networks (P-Prob < 0.50). On the bottom, P-Prob have been fixed to 0. The more decreases the R genes concentration, the more distributed become the network.

If there is R this means that the receivers of a copy of a genome are not allowed to resend the replicated plasmid to another agent. It avoids decentralized propagation of strategies, considering that the original owners of a genome are the only ones that can distribute copies. High possibility of R implies a constraint to diffusion of received strategies, because receiver will be able to use the successful strategy but will not be allowed to propagate them and share his knowledge with others.

Decentralization is inversely related with these two parameters. High P and R rates imply centralized societies without *P2P* communication and without reproduction rights. Oppositely, low P and R rates lead to *P2P* exchanges of information without limits of copies.

During a conjugation process, when one agent sends a plasmid to another, the S sequence could be modified. This means that mutation of any strategy is allowed by default. But mutation can be inhibited if O is present in the genome. O sequence implies that a plasmid cannot be modified. So only low levels of O presence lead to an open society in which variation of bad strategies in short time is guarantee. However, high presence of O in the population genome implies that strategies are closed and invariant. So once an agent follows a specific strategy he cannot change it until he receives another genome from a more successful agent.

Differentiation of strategies is another important variable in this model. Cardinality of Sec is related with the number of different strategies by default; so if |Sec| is near to |A| and there is a low presence of O segments in population genome, then it implies more heterogeneity.

Following that approach, our agents behave according to their own internal states by following single algorithmic rules. Each agent's decision takes into account two values, his own knowledge and his neighbor's knowledge. Each iteration agents compare both parameters by using their *internal evaluator,* which is the function that constitutes their bounded rationality. There is a different *internal evaluator* to every single agent, assuming variability of cognitive skills within a population. After positive evaluation, if it is worth to learn a new strategy according to their criteria (and if any inhibitor impede it), conjugation between two agents occurs. When an agent receives a new genome it replaces the previous one. This replacement can be complete or not, depending of mutation inhibition. In this version of our model, mutation implies a recombination of 50% of both genomes. After replacement of genome, the receiver tests his new strategy by using it to optimize a fixed *selection function.* Accuracy of function optimization determines the new knowledge level of the agent. The more knowledge an agent achieves, the more social reputation he obtains.

## 3 Experimental Results

In order to analyze the influence of specific variations on initial conditions, we have simulated different scenarios. The focus of this study is how probabilistic distribution of Boolean genes (P, R, O) affects social structure, that is, the role of cultural constraints. We have fixed |Sec| ~ |A| and recombination to 0.5 when mutation is allowed. For any of those setup configurations, we have executed our model during 50 iterations. Repeating each one of these experiments with random populations of $10^4$ agents we have observed common patterns that are related with P, R and O presence in population genome. We have tested the emergence of different global configurations, from centralized societies with low levels and unequal distribution of knowledge to "P2P societies" in which heterogeneity and decentralization lead to collective success. We have used NetLogo 5.0.4 to implement our model and R 3.1.1 to plot the simulation data.

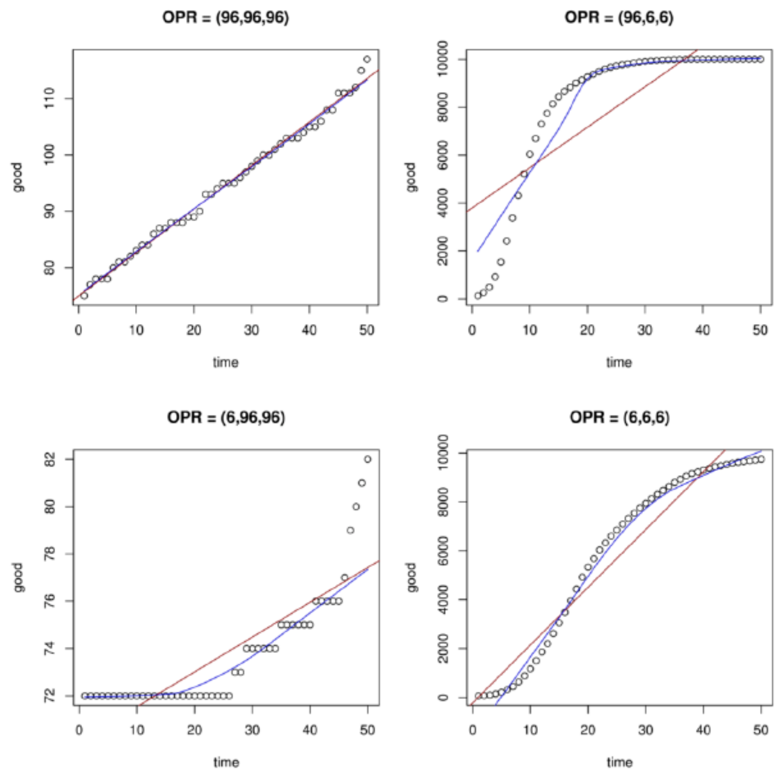


Figure 4: Impact of OPR on production of knowledge. Plot of strategies with accuracy higher than 0.7 in four simulations of $10^4$ agents during 50 iterations.

We have defined a selection function with ten variables. Every turn, agents replace each one of those variables with the values of the S segment. If these ten genes optimize the function and the result is equal or greater than 70, we consider that the owner of that $S_i$ segment has a good level of knowledge. The adaptive behavior of the system as a whole consists in increasing the distribution of good strategies and the elimination of bad ones. As we can see in the figure 4, performance is mainly related with P presence, that is, with the number of agents that are allowed to share strategies.

After reproducing several scenarios, we have observed that worse results are in centralized and homogeneous societies; for example, a society with presence of O and R in the whole of population genome. O presence means that we will not see strategy modifications, that is, the maximum level of knowledge will be static during the $10^4$ iterations. Then the best strategy will depend only on initial configuration, when random $S_i$ segments are generated; this simulation will only take into account the propagation of those initial strategies. R implies that there is not any reproduction of received genome, so only the original owner will be able to propagate it.

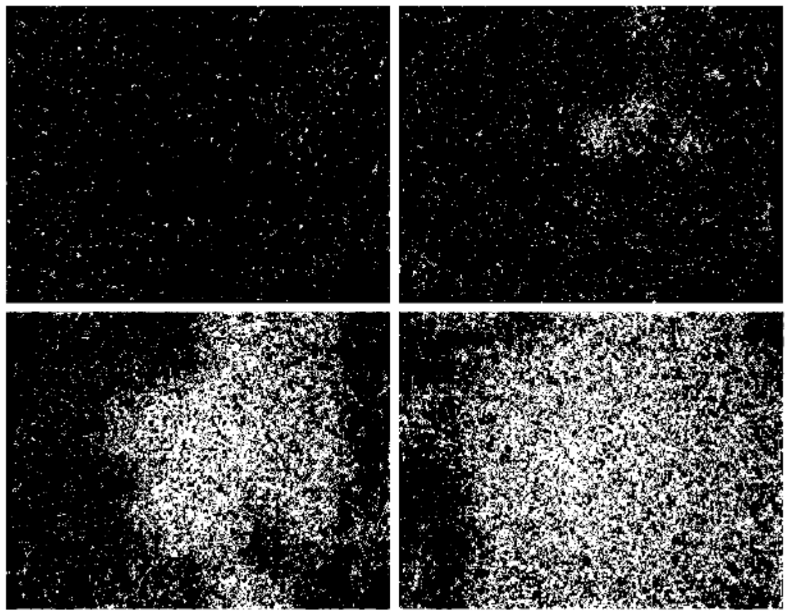

Figure 5: Distribution of knowledge in a "P2P Society". Low P, O and R rates in a population with $10^4$ agents. Peer-to-peer exchanges, reproduction and modification rights lead to this map of successful strategies. Each picture represents a snapshot of the same simulation in time.

We have studied how P presence modifies the global behavior of this system. In order to compare two variations of that artificial society, one with P probability ~ 0.90 and other with P probability ~ 0.19, we have repeated the experiment several times, testing that centralization leads to low results in knowledge generation. We have also simulated scenarios with high heterogeneity and decentralization levels by reducing P, R and O probabilities and we have seen that with a whole elimination of O sequences (O *probability ~ 0*), that is, activating mutation of strategies, generation of knowledge not only is faster but also richer in variety. As picture 5 shows, different focus on improvement are found in population when diversity occurs. In this heterogeneous scenario, good strategies come from different agents with different initial locations. Agents evolve in creative ways because of modification of strategies is allowed and innovative knowledge is propagated because of decentralization, leading to an egalitarian "*Knowledge Society*". Another example of how heterogeneity is related with innovation can be found in the figure 6, according to the results of an extended version of the basic model. This model implements a sugarscape-like scenario with dynamic distribution of resources and without a selection function. In this case, the strategic genome (S) codifies the motor behavior. Low O-Prob implies more heterogeneity of strategies and therefore a resilient behavior.

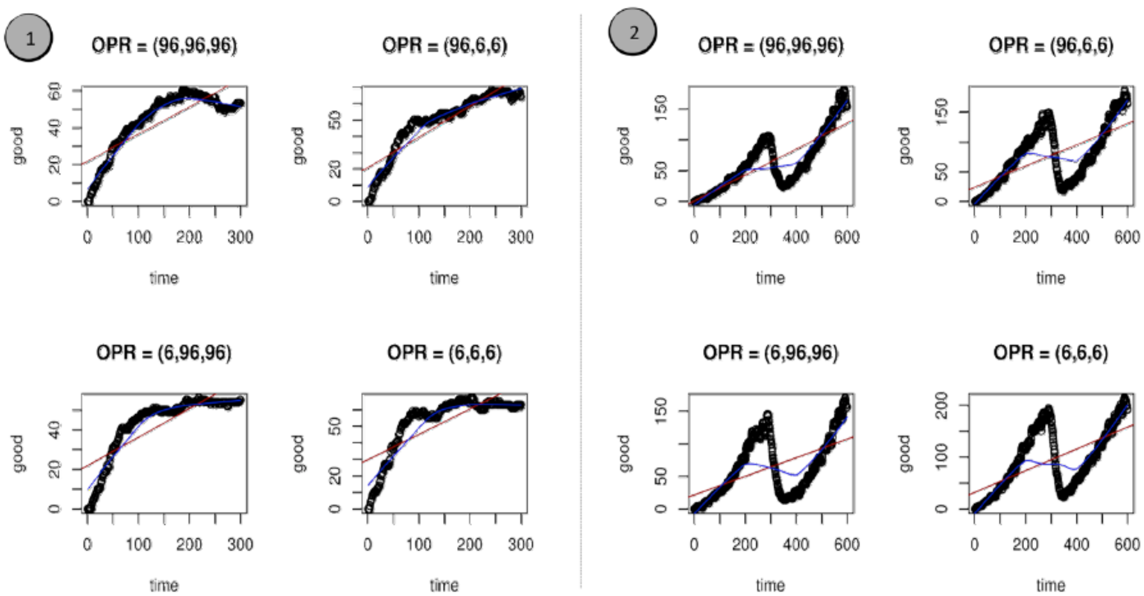


Figure 6. Impact of OPR on adaptation to a dynamic environment with a swift change in resources distribution at time = $3 \cdot 10^2$. Plot of strategies with accuracy higher than 0.7 in several simulations of $10^2$ agents during $3 \cdot 10^2$ iterations (1) and $6 \cdot 10^2$ iterations (2). Decentralized systems perform better because they allow horizontal gene transference, that is, horizontal learning. Heterogeneity produces more innovative solutions preserving *nomadism* in *sedentary* communities.

## 4 Conclusions

We have modeled complex social dynamics with bacterial-based algorithms. We have seen how in our artificial society, system optimization is limited by centralization and homogeneity. Based on experimental results, we can infer that egalitarian distribution of optimal strategic knowledge can be achieved by changing some cultural values in population, but also technical and political restrictions that constrain access, reproduction or modification of information. Considering that strategic optimization constitutes the wealth of this model as much as actual knowledge constitutes the main good of social systems at the Internet Age, we conclude that human societies will achieve their optimal configuration only by incrementing decentralization and heterogeneity.

Humanity can achieve social development in an egalitarian way by producing distributed and open knowledge through peer-to-peer exchanges in heterogeneous communities connected by digital networks. The development of a sharing ethic built on decentralization and heterogeneity has been the basis of projects like Wikipedia or GNU/Linux. In order to achieve an efficient and distributed generation of knowledge, it is necessary to keep walking in the same direction. Information policies, academic institutions and e-learning approaches have to embrace public domain licenses, free software and open data. By working in that way, it will be possible to continue learning together, growing global communities, generating innovation and building what we have named a "P2P Society".

## References

Barabási, A.-L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature reviews. Genetics*, *5*, 101–113. doi:10.1038/nrg1272

Baran, P. (1964). On distributed communications: Introduction to distributed communications networks. *Volumes IXI RAND Corporation Research Documents August*, *12*, 51.

Bauwens, M. (2005). The political economy of peer production. *Retrieved on Nov*, *23*, 2008.

Bauwens, M., Sussan, R., Evolution, H., Summary, E., Sociology, T., Framework, T., & Capitalism, C. (2005). Peer to Peer and Human Evolution, 1–190.

Cioffi-Revilla, C., & Rouleau, M. (2010). MASON RebeLand: An Agent-Based Model of Politics, Environment, and Insurgency. *International Studies Review*, *12*(1), 31–52. doi:10.1111/j.1468-2486.2009.00911.x

Deffuant, G., & Gilbert, N. (Eds.). (2011). *Viability and Resilience of Complex Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-20423-4

Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies*. *Proceedings of the fifth annual CCSC northeastern* (p. 208).

González Rodríguez, D. (2011). Simulador de bacterias sintéticas: la conjugación como protocolo de comunicación entre células. *I Jornadas Doctorales de Castilla-La Mancha (Resúmenes de Comunicaciones).*

Heylighen, F. (1989). Self-organization, emergence and the architecture of complexity. *Proceedings of the 1st European conference of Complexity.*

Lansing, J. S. (2003). Complex Adaptive Systems. *Annual Review of Anthropology*, *32*, 183–204. doi:10.1146/annurev.anthro.32.061002.093440

Mitchell, M. (1999). An introduction to genetic algorithms. *Cambridge, Massachusetts.*

Tesfatsion, L. (2003). Agent-based computational economics: modeling economies as complex adaptive systems. *Information Sciences*, *149*(4), 262–268. doi:10.1016/S0020-0255(02)00280-3